Feature Set Evaluation for Classifiers
Progress Report # 1

November 20, 1988

period: September 15, 1988 to November 15, 1988
contract # : DAAB07 – 88 – C – F415
security classification : unclassified
CDRL # : A001

for: Director
Center for Night Vision & Electro Optics
ATTN : AMSEL – RD – NV – AC
(Martin Lahart)
Fort Belvoir , VA 22060 – 5677

by: KAB LABORATORIES INC.
3116 Mercer Lane
San Diego, CA 92122
(619) 455 – 5361

TABLE OF CONTENTS

	page
I. INTRODUCTION	
A. Background	1
B. Objectives	1
C. Scope	3
II. RESULTS	3
III. STATUS	7
IV. PLANS	7
V. CONCLUSIONS	10

I. INTRODUCTION

A. Background

On September 13, 1988 KAB LABORATORIES INC. (KABL) was awarded a Small Business Innovation Research (SBIR), Phase I contract with the Center for Night Vision & Electro Optics (CNVEO). Work on the contract commenced on September 15, 1988. The Phase I activity is to conduct research aimed at developing techniques for improving automatic recognition/classification of targets. This is the first progress report under that activity, and covers the first two months of a six month contract.

B. Objectives

Automatic Target Recognizers (ATRs) have tried a wide variety of feature set classifiers in attempting to improve the quality of their classification of targets. The selection of these feature set classifiers to date has largely been based upon subjective intuition of the analyst. The analyst typically approaches the problem by starting with a proposed feature set which is derived somewhat heuristically based on an analyst's understanding of the underlying physical phenomena which differentiate a target from any background "clutter" or "noise" which may exist. This underlying phenomenology can be exceedingly complex in the case of real military targets, in real clutter filled backgrounds, imaged by electro-optical sensors under the less-than-ideal circumstances which may exist in a battle field environment.

The feature set for ATR applications could easily contain a large number of individual features or measurements (e.g., location of hot spots, geometric ratios, areas, perimeters, texture mixture, etc.). For real time systems, these features must be extracted quickly and processed to determine the target identification (classification). To minimize computations and keep ATR

processor requirements at a reasonable level, the ATR algorithms should be efficient and extract only those features which are most useful to the identification process. The selection of this set of reduced features which possess the most powerful discriminating capability is the subject of this study.

KABL has proposed to use an existing software package, developed by PAR Government Systems Corporation (PGSC), called the On-Line Pattern Analysis and Recognition System (OLPARS) as a tool for feature set analysis. Under this contract CNVEO will be furnished with an OLPARS licence, software, and documentation. The OLPARS will also be enhanced by our research to include a new promising feature set evaluation algorithm aimed at meeting specific CNVEO needs.

The Phase I SBIR activity has proposed meeting the following five technical objectives:

1. identify and propose a collection of feature set evaluation algorithmic tools which address unique characteristics of feature sets used in ATR applications.
2. implement at least one new promising feature set evaluation algorithm in FORTRAN and integrate it within the On-Line Pattern Analysis and Recognition System (OLPARS), which is an existing commercial software system which provides general purpose feature set evaluation and classifier design capabilities.
3. demonstrate the performance of the new feature set evaluation algorithms already within OLPARS using feature sets derived form both real and simulated E/O imagery.
4. provide DoD with a licenced VAX-compatible copy of the augmented OLPARS software package.
5. document the proposed new set of feature set evaluation algorithms and the test results obtained with the newly implemented algorithm within a final technical report.

By using the OLPARS in our research we will be taking advantage of considerable previous work on this subject. The OLPARS was initially developed in the early 1970's as a pattern analysis support tool. Since that time it has been enhanced to increase its capability for analysis and display and to make it user friendly. It also comes with full supporting documentation. Upon completion of the Phase I activities CNVEO would possess an independent capability to analyze, select and test feature sets and to evaluate their relative discriminating power for target classification. This capability should provide a means for both improving and testing their own ATR approaches and for evaluating the approaches suggested by industry.

C. Scope

This report covers the first two months of a six month study. The Phase I activity calls for $25,000 of material cost for the purchase of OLPARS, computer time, and a subcontract to PGSC for 75 man-hours of support on the OLPARS program. The remaining $25,000 is spread over 6 months for KABL manpower to support research, and for incidental costs such as travel.

In the sections which follow the progress and plans of the activity will be reported. Section II. RESULTS, will present the work to date and accomplishments. Section III. STATUS, will summarize manpower expenditures to date and relationships to program milestones. Section IV. PLANS, will present the major planned activities for the next four months. Finally, Section V. CONCLUSIONS, will summarize the findings to date.

II. RESULTS

The program schedule given in the SBIR proposal called for completing Task 1 (Feature Set Evaluation) and starting on Task 2 (Algorithm Implementation) in the first two months. KABL has

kept to that schedule.

Task 1 included the following five sub-tasks:

1. meet with the sponsor to discuss specific goals for the effort, possible algorithms, test data sources and planned use of OLPARS;
2. acquire OLPARS licence;
3. investigate, propose, and document feature evaluation algorithm set;
4. perform tradeoff analysis to rank the individual algorithms based on:
   - expected performance/payoff
   - complexity/ease of implementation
   - applicability to problems of primary Army interest;
5. select a candidate algorithm for implementation within the OLPARS framework.

The sponsor for this work, Mr. Martin Lahart, visited San Diego in late September. We took advantage of this opportunity to give Mr. Lahart a brief tutorial on OLPARS and a demonstration of the system. We also obtained further detail on CNVEO's primary areas of interest. Armed with this information we obtained and reviewed a number of research papers pertaining to their work. This research, carried out on reports from Mr. Lahart, from the Naval Ocean System Center library, and from the University of California San Diego libraries enabled us to focus on the primary needs of the CNVEO.

A second meeting with Mr. Lahart was held on October 27, 1988 at the CNVEO, Fort Belvoir, VA. The principal investigator, John Konotchick, and a PGSC representative, David Robbins were in attendance. At CNVEO request, Mr. Robbins presented an overview briefing of the OLPARS to a number of Center personnel. Following the briefing, Mr. Lahart provided us with a description of CNVEO equipment we might interface with, and also a list of

the key areas of OLPARS enhancement of most interest to CNVEO. Our purpose in the visit was to be responsive to the desires of CNVEO and so this list, rather than our own would be used to select a feature set evaluation algorithm for development. The list included six possible enhancements, as follows:

1. Computation of error rates using assumed distribution and error rates;
2. Geometric transformation - How are features and error rates changed?;
3. Identifier for particular points in feature space - Mechanize an interface with ORACLE;
4. Provide a four-dimensional display of a form discussed at CNVEO;
5. Analyze relative discrimination ability of pairs of features;
6. Provide a metafile for plotting - to generate hard copy and displays.

These six possible enhancements had been discussed either in the OLPARS briefing meeting, or privately with Mr. Lahart, and were commonly understood by the KABL team and Mr. Lahart. We were to study these and report back on which, if any, could be implemented during Phase I.

After considerable discussion and analysis by PGSC and KABL it was decided to attempt to implement #5. on the list. Algorithms analyzing pairs of discriminators had never been tried on OLPARS, but it was felt that it would add a powerful addition to the planned CNVEO capability.

The OLPARS system provides a number of discriminants for ranking an individual feature's ability to discriminate a class from all others, or ranking a feature's ability to discriminate between two classes. It does not, however, have the ability to rank "pairs" of features for their ability to discriminate classes.

The enhancement which will be attempted under the Phase I research effort will provide this capability to the CNVEO system.

A critical measure of the ability for feature pairs to discriminate classes is their probability of misclassification. The exact calculation of this error is often impractical or impossible, however, and so other related measures are often chosen. The most common approach is to define a separability measure, or distance, between the probability distributions of the classes under investigation. Assuming that the most important characteristic of this distance measure is its upper bound on error (of misclassification), we can rank feature pairs by their ability to minimize this error. This implies a distance measure with a known relationship to an error upper bound. A number of distance measures for these feature pairs have been derived (e.g., Matusita's, Vajda's entropy, Devijver's Bayesian distance, Ito's measure, Komogorov's variational, Toussaint's, etc.), but the Bhattacharyya distance is one that both provides a reliable measure, and one which could be easily implemented on the OLPARS.

The Bhattacharyya distance will provide a measure of which pairs of features have the highest separability between classes. All possible feature pairs can then be examined to determine their relative ability to discriminate between all possible class pairs. The Bhattacharyya measure which will be analyzed and implemented will assume a Gaussian class distribution.

The original Phase I schedule called for the enhanced OLPARS to be delivered to CNVEO at the end of Phase I. After our visit on October 27th, however, we were asked if the basic OLPARS could be provided as soon as possible to CNVEO. KABL discussed this with PGSC, and received their approval to install OLPARS in the week of November 14-18, 1988. The quick reaction response of PGSC is the more laudatory because they scheduled the installation before

either preparing the licence agreement for CNVEO or the invoice for KABL.

This early OLPARS delivery, while causing minor schedule and plans changes, is not expected to affect major schedule milestones.

III. STATUS

While meeting schedule goals the KABL team has remained flexible and responsive to CNVEO specific desires. As the previous section has described, we have undertaken an ambitious challenge within our manpower allotment of the Phase I program. We have kept to the schedule of tasks proposed in the SBIR proposal, and have only slightly exceeded the manpower projections. Figure 1 below presents a plot of current manpower expenditures for the first two months, overlaid on the total Phase I allotment. Table 1 presents the data in tabular form.

Figure 2., reflects the new schedule as agreed to by CNVEO and KABL. There are three minor changes in this new schedule as compared with the initial proposed schedule. These are: 1. the Task 2 and Task 3 activities have been extended 2 weeks (to reflect our desire to implement a CNVEO option); 2. the OLPARS delivery has been pushed forward from February 1989 to November 1988 (to reflect CNVEO desire to use the system early); and 3. the number of meetings with CNVEO has increased from the planned 2 to 4. These changes are not expected to significantly affect the overall Phase I research effort.

IV. PLANS

The major near term activity of the KABL team will be to develop

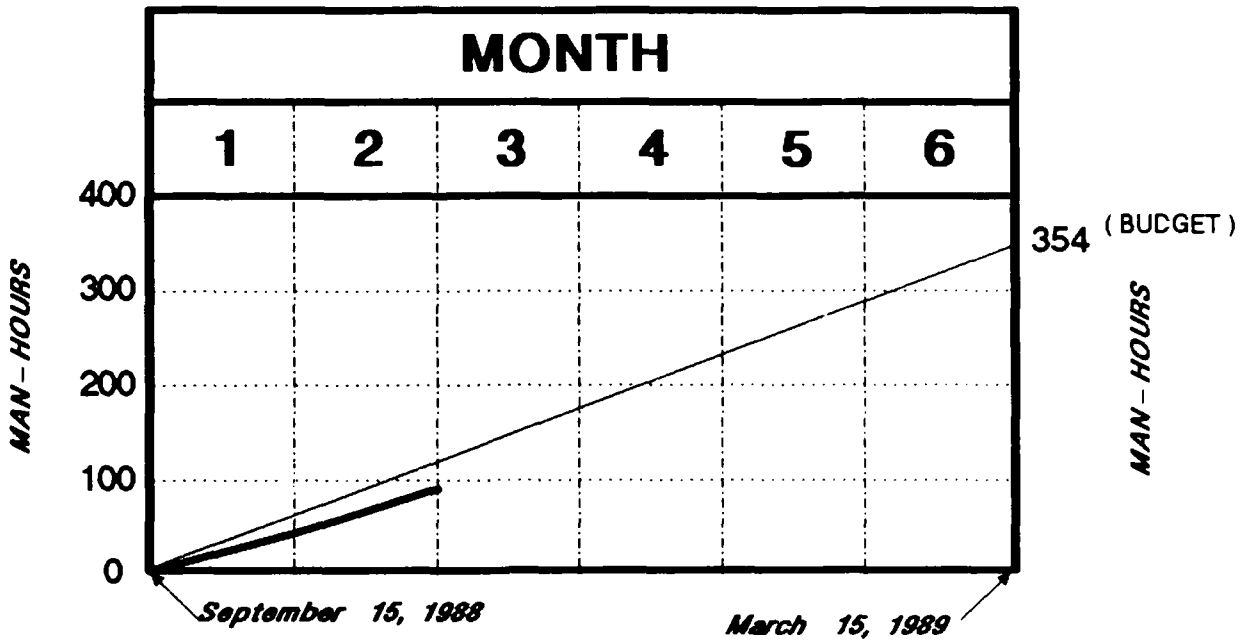

Figure 1. MAN-HOURS DELIVERED

PERSONNEL BUDGET/ACTUAL		MONTH 1	2	3	4	5	6
KABL	principal investigator	36/40	36/45				
	research assistant	6/0	6/0				
PGSC	senior staff member	15/0	15/0				
CUM. TOTAL		57/40	114/85				

Table 1. MAN-HOURS DELIVERED

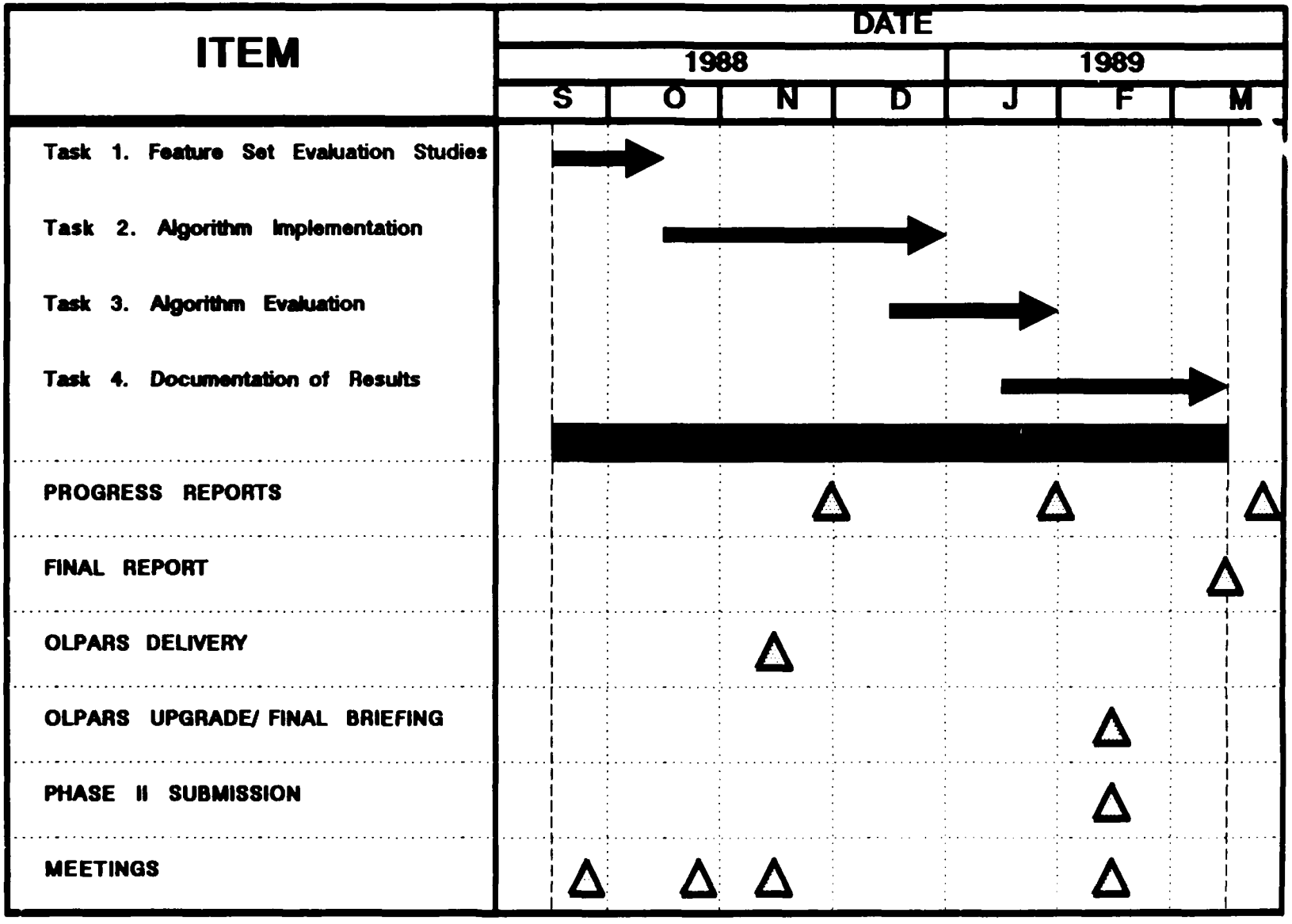

Figure 2. PHASE I SCHEDULE

the evaluation algorithm (Bhattacharyya distance) described in Section II, which will provide measures of the discriminating ability of pairs of features. This will involve analytical analysis of the mathematics involved, analysis of the OLPARS coding structure, and coding of the technique. Following coding on the PGSC, in-house OLPARS, it will be tested and de-bugged. When it has proven to be effective it will be presented to CNVEO at the final briefing meeting and installed on their OLPARS. Finally, it will be documented in the final technical report of Phase I.

During the remaining months of Phase I, the KABL team will also plan to perform initial analysis of data sets provided by CNVEO. It is expected that close contact with the CNVEO sponsor, Mr. Lahart, will be maintained throughout these final four months. With the early delivery of OLPARS it is also expected that some tutorial consulting on OLPARS operation may be required.

V. CONCLUSIONS

The KABL team is focusing its research on specific desires of CNVEO. The initial two months activity were highly successful, due in large part to a universally cooperative spirit between all participants in industry and at CNVEO. The willingness of CNVEO personnel to track down details of their equipment for us, to furnish research papers, sample data sets, and to spend the time describing their work and goals has helped greatly in getting this research effort quickly off-the-ground. The willingness of KABL's subcontractor, PGSC, to expedite OLPARS delivery schedules and share their pattern recognition experience have also been greatly appreciated. Although, the development of the planned 2-feature evaluation algorithm remains a challenge, especially within the Phase I manpower constraints, no major problem areas are apparent. The progress, to date, has been as planned.